#48, 80

RESULT 7
AAR40841
ID AAR40841 standard; Protein; 3054 AA.
XX
AC AAR40841;
XX
DT 02-MAR-1994 (first entry)
XX
DE Translation of TEV large ORF.
XX
KW Tobacco Etch Virus; TEV; large nuclear inclusion protein;
KW capsid protein; vector; transformation; Agrobacterium tumefaciens;
KW potyvirus.
XX
OS Tobacco etch virus.
XX
FH Key   Location/Qualifiers
FT Protein   2791..3054
FT   /note= "TEV coat protein"
XX
PN WO9317098-A.
XX
PD 02-SEP-1993.
XX
PF 18-FEB-1993; 93WO-US01544.
XX
PR 19-FEB-1992; 92US-0838509.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Dougherty WG, Lindbo JA;
XX
DR WPI; 1993-288404/36.
DR N-PSDB; AAQ47788.
XX
PT Plant transformation vector - comprising Tobacco Etch
PT Virus-derived mutated gene, conferring viral resistance to e.g.
PT tobacco plants infected with poty-viruses
XX
PS Disclosure; Fig 1; 68pp; English.
XX
CC This sequence is encoded by the largest ORF in the third reading
CC frame of the Tobacco Etch Virus (TEV). This protein may be cleaved
CC to give the large (54 kD) nuclear inclusion protein and the capsid
CC protein. The DNA encoding this protein was used as a basis in the
CC construction of a vector suitable for transformation into Agrobacterium
CC tumefaciens (AT) cells. The vector constructs of the invention (see
CC also AAQ55665-67) encode a full length coat protein and a truncated coat
CC protein (see AAR45612-13). The resultant transformed AT cells may
CC be used to transform cultured tobacco cells. The transformed cells
CC had reduced susceptibility to viral infection, ie. they were resistant
CC to potyvirus infection.
XX
SQ Sequence 3054 AA;

  Query Match   47.6%; Score 1085.5; DB 14; Length 3054;
  Best Local Similarity   49.3%; Pred. No. 5.3e-103;

```
  Matches  212;  Conservative   76;  Mismatches  137;  Indels    5;  Gaps    4;

Qy       1 GKSKRTRQKLKFRAARDMKDRYEVHADEGTLVENFGTRYSKKGKTKGTVVGLGAKTRRFT 60
           ||  |   :  |||  |  ||   :  :|||  |:    |    ||:  |:  |||  |||   |:|||:|:|
Db    1850 GK-KNQKHKLKMREARGARGQYEVAAEPEALEHYFGSAYNNKGKRKGTTRGMGAKSRKFI 1908

Qy      61 NMYGFDPTEYSFARYLDPITGATLDETPIHNVNLVAEHFGDIRLDMVDKELLDKQHLYLK 120
           ||||||||::|:  |::||:||  |:||:     ::||    ||  :|   |:   :  ::  |  |
Db    1909 NMYGFDPTDFSYIRFVDPLTGHTIDESTNAPIDLVQHEFGKVRTRMLIDDEIEPQSLSTH 1968

Qy     121 RPIECYFVKDAGQKVMRIDLTPHNPLLASDVSTTIMGYPEREGELRQTGKARLVDPSELP 180
             |  | |     :||:::|||||: | ||:  ||  |||:|||| ||||||  |   |    :||
Db    1969 TTIHAYLVNSGTKKVLKVDLTPHSSLRASEKSTAIMGFPERENELRQTGMAVPVAYDQLP 2028

Qy     181 ARNEDIDAEFESLNRISGLRDYNPISQNVCLLTNESEGHREKMFGIGYGSVIITNQHLFR 240
           :|||:   |  |||     | |||||||   :|  |||||:||    ::|||:|   ||||:||||
Db    2029 PKNEDLTFEGESL--FKGPRDYNPISSTICHLTNESDGHTTSLYGIGFGPFIITNKHLFR 2086

Qy     241 RNNGELSIQSKHGYFRCRNTTSLKMLPLEGHDILLIQLPRDFPVFPQKIRFREPRVDDKI 300
           ||||  |  :||  ||  |:  :|||:|:    ::|  |:::|::|:|||  ||||::||||:  :::|
Db    2087 RNNGTLLVQSLHGVFKVKNTTTLQQHLIDGRDMIIIRMPKDFPPFPQKLKFREPQREERI 2146

Qy     301 VLVSTNFQEKSSSSTVSESSNISRVQSANFYKHWISTVAGHCGNPMVSTKDGFIVGIHSL 360
           ||:||||  ||  ||  ||::|          |:||||  |   |  ||:|:|:|||:|||||||||
Db    2147 CLVTTNFQTKSMSSMVSDTSCTFPSSDGIFWKHWIQTKDGQCGSPLVSTRDGFIVGIHSA 2206

Qy     361 ASLTGDVNIFTSFPPQFENKYLQKLSEHTWCSGWKLNLGKISWGGINIVEDAPEEPFITS 420
           ::  |    |  |||  |   |  :  |         | |||:||   :  |||  :    |||||
Db    2207 SNFTNTNNYFTSVPKNF-MELLTNQEAQQWVSGWRLNADSVLWGGHKVFMSKPEEPFQPV 2265

Qy     421 KMAS-LLSDL 429
           |  |:  |:::|
Db    2266 KEATQLMNEL 2275
```